APPROVED FOR RELEASE 1994
CIA HISTORICAL REVIEW PROGRAM
18 SEPT 95

TITLE: R&D For Intelligence Processing

AUTHOR: CODIB Task Team VI

VOLUME: 10 ISSUE: Spring YEAR: 1966

STUDIES IN INTELLIGENCE

A collection of articles on the historical, operational, doctrinal, and theoretical aspects of intelligence.

All statements of fact, opinion or analysis expressed in Studies in Intelligence are those of the authors. They do not necessarily reflect official positions or views of the Central Intelligence Agency or any other US Government entity, past or present. Nothing in the contents should be construed as asserting or implying US Government endorsement of an article's factual statements and interpretations.

Recommendations for invigorating and coordinating the community's development of data-handling systems.

R&D FOR INTELLIGENCE PROCESSING

CODIB Task Team VI[1]

The team considered conceptual and managerial aspects of establishing R&D programs for intelligence data handling to be more crucial and more in need of immediate attention than technical aspects. Rather than concern itself with what technical approaches should be adopted, what type of equipment is best suited to a particular application, and the like, it therefore sought answers to such questions as the following. To what degree are the several USIB members' R&D programs for intelligence data handling mutually supportive? Are existing and planned programs adequate in size, balanced in content, technically sound, and adequately organized, managed, and funded? Can the technical leadership for such programs be improved? How should policy be established for a coordinated community program? What outstanding opportunities might be seized as immediate practical objectives of R&D? How might shortcomings in present data handling be translated into R&D requirements and communicated to the technical leadership of the community?

Community Objectives

The team set out to find a framework of R&D goals with respect to data handling in the intelligence community to which to relate specific managerial and technical tasks and within which to identify deficiencies and achievements. It discovered instead that the community, as governed by USIB under established intelligence directives, has no organized set of R&D objectives (except as NSA and NPIC are in-

[1] Adapted from portions of its report dated 28 September 1965. The team, charged in the preceding March with defining interagency goals for R&D in the processing of intelligence data, had representation from all USIB agencies except the FBI and AEC and the help of consultants provided by the National Science Foundation. It was chaired by Dr. Ruth M. Davis of the Office of the Secretary of Defense. This adaption does not necessarily reflect the views of USIB or its Committee on Documentation; neither body has yet completed action on the report.

dividually charged with R&D for their respective specialized purposes), no policy for establishing objectives, and no mechanism for either. Although the federal government as a whole has similarly no explicitly stated R&D goals, there is a formal mechanism within the executive branch for advising the President on R&D, coordinating agency programs, and picking particular areas for concentrated attention—the Scientific Advisor to the President, his staff in the Office of Science and Technology, and the panels and committees over which he presides. The USIB community is represented in this mechanism only insofar as its member agencies are individually represented.

The intelligence community in many ways functions as a self-contained entity isolated from the rest of the federal structure by organizational, managerial, and security barriers. This isolation causes it little or no distress in operational affairs, but scientific and technical activities are another matter. Here the community is not self-sufficient. These activities, not only managerial and R&D but testing, engineering, evaluation, and implementation, are often delegated in part or whole to groups outside the community.

Much such delegation of R&D on data handling is useful, often essential. The bulk of technical competence in general information handling—information sciences technology—lies outside the intelligence community, and many aspects of intelligence data handling are identical to those of general information handling. Multi-font optical readers, for example, new storage media, large random-access memories, automatic translation, and improved man-machine communications are needed equally in intelligence and outside. There is no reason why USIB agencies should bear a disproportionate share of such R&D costs. Sharing these with others should permit the intelligence community to concentrate its limited resources on those data-handling needs which are of unique or primary concern to its missions, not only those of NSA and NPIC but also the indications and warning mission and many others.

Thus the concentration outside the intelligence community of technical competence with respect to intelligence data handling may be viewed without alarm. Lack of competence within the community in the applications of data-handling techniques to intelligence problems or systems is unjustified, however, and the team believes that at present such competence is marginal at best. This belief is backed by the frequent use of contractors for system design and development, the mediocrity of data-handling techniques and systems currently

used in the community, and the apparent lack of concrete planning for the application of more sophisticated technology.

It is frequently asserted that a lack of federal or national objectives can be compensated for by well-structured and documented individual agency objectives, so it may be that well-founded USIB member agency R&D objectives in intelligence data handling would provide a suitable substitute for the lacking community goals. The task team attempted, then, to discover individual agency objectives in order to assess their suitability. It was found that DIA, the military departments of DoD, and NSA had documented objectives. The State Department had none. The existence of CIA objectives was not determined, and the NSA objectives were not released to the team. It was obvious, however, that the objectives identified were not uniform in structure, were neither comprehensive nor cohesive, were grossly incomplete with respect to managerial considerations, and were not intended as guidelines for R&D efforts. It was necessary to conclude that the aggregate of agency objectives could not substitute for USIB objectives and that the individual objectives were of little use in judging or interrelating R&D efforts planned or under way.

Recommendations

In view of the importance of R&D in intelligence data handling, the need for sharing responsibility for it with groups outside the USIB community, the absence of any USIB goals, policies, or mechanism to further the necessary R&D, and the lack of coordination among present and planned R&D efforts, the task team recommended a set of actions. These actions are aimed at improving the managerial position of the community not only in handling its internal operational requirements but in its dealings with other federal offices and groups outside the government. It appears axiomatic that both the community and its member agencies will benefit in their individual and joint contacts with outside entities if they can assume a uniform and professional negotiating posture. The actions recommended are discussed below.

Policy mechanism. The USIB should set up a permanent body to establish community objectives and policies for R&D in intelligence data handling. This body should have representation from all member agencies and a full-time executive secretary. It should coordinate the agencies' objectives, plans, policies, and evaluations and be the principal advisor to Chairman, USIB, in this field. In critical R&D areas it might, if deemed advisable, develop its own plans and rec-

ommend where the R&D responsibility should be delegated. It should have authority to get information on the agencies' R&D planning and budgeting but none to approve plans. Since its membership could not have the technical competence necessary to cover the whole field of intelligence data handling, it should have continuous access to consultants both within and outside the community. It should meet at least monthly.

Reporting service. The USIB should establish a formal mechanism to disseminate technical information concerning current and planned data-handling R&D in the community. This service would handle only classified information, and it would utilize channels affording the needed security. It would be given access to information in the agencies, select that to be disseminated, and give it maximum sanitization before dissemination. It has been found that linking technology to the sponsoring organization or to the operational use for which it is intended normally increases the security classification of a document and consequently the inaccessibility of the technical data in it.

For open-source materials, existing information services appear to be adequate. If it is found that they are not, additional requirements can be levied on them so that the USIB reporting system does not have to handle such easily accessible information.

The task team believes that the classified information can be drawn from existing agency reporting mechanisms functioning with little change. Bleed-off from these established systems should suffice initially, at least, and should serve to point up new requirements, if any, which would have to be imposed on them. If judicious use is made of individual agency personnel through the USIB policy body, this reporting service should need funds only for administrative support—clerical and mail services, etc.

Stimulation of personnel. Those responsible for R&D in intelligence data handling need to be impressed with the importance of keeping up with others' pertinent research and development. IDH/R&D personnel should be vigorously encouraged, if not gently coerced, to make use of the open-source information services available to them as government employees. They should be subjected to scheduled evaluation to measure their effectiveness, competence, and awareness of current R&D. The importance of their functions in terms of the responsiveness of the intelligence community to any situation, crisis or normal, cannot be overestimated; and yet they exist as an unstructured, unrecognized, and uncoordinated group with no group allegiances and no reward-punishment mechanism.

The task team had extreme difficulty in even identifying those responsible for R&D in intelligence data handling and certainly found no IDH/R&D community. Many who had responsibility for a data-handling project were not even aware of anyone else having similar responsibilities. The policy body recommended above should establish an agreed-upon organizational listing of IDH/R&D components and a directory of IDH/R&D personnel giving their individual specialties.

IDH/R&D technical personnel presumably do not differ generally in work habits from other government technical personnel. A DoD study of the information usage habits of government scientists and engineers made last spring should accordingly be applicable to them, and no separate survey of them should be needed. This DoD study, along with other evidence, points to either misuse or inadequate use of information services by technical personnel and attributes it primarily to lack of instruction. The team's recommendation, therefore, is that USIB arrange for the compilation of a report listing the 400-500 available information services and giving details on their accessibility and procedures for their use. Twelve months after distribution of the report, a study should be made to evaluate changes in information usage patterns brought about by it. This could then be the basis for recommendations for improvements. These measures—a directory of IDH/R&D personnel, a report on available information services, and a follow-up study of usage patterns—could be accomplished by a contractor under USIB supervision.

Feedback system. Feedback from users of intelligence is not systematized, nor is the extent of feedback and its impact known. The mechanisms now existing—post mortems, validity studies, field comment, consumer comment—provide limited return and this largely confined to National Intelligence Estimates. There has been little contact between intelligence analysts and IDH/R&D personnel.

It would thus appear that the nature, level and extent of feedback should be studied and the feasibility of more systematic dialogue between producer and consumer at various levels explored. The study would require the services of personnel particularly talented in the production process to work with experts in techniques for evaluating output.

The Price of Inaction

If these recommendations are not accepted and some such line of action taken, the intelligence community will continue vulnerable to

external investigative and evaluative groups, with no recognized negotiating position from which to meet questions concerning intelligence data handling. The field of information sciences and services is a highly populated one in the scientific community. The product of the intelligence community is information, and intelligence data handling is analogous to information sciences technology. So one can expect a high outside interest in the intelligence community's R&D in data handling. Such interest is good and should be maintained; and investigations can be extremely productive if a comprehensive picture is presented to the investigators. But this has often not been the case in the past.

It is elementary management doctrine that when there is a responsible coordinating mechanism deficiencies are fewer and those which do exist are easier both to find and to correct. It would help both the external investigators and the community to have such a coordinating mechanism. Regardless of how well intentioned an investigating group may be, when the data presented it are fragmentary and not interrelated, its recommendations are even more fragmentary and unrelated to the real problems. They may "remove a thorn and by so doing implant a tumor." The intelligence community can ill afford any more such investigations.

Internally, the deleterious effects of having no over-all interagency or USIB objectives or policy in data-handling R&D are felt in every phase of the R&D activity. There is no structure on which to hang R&D efforts other than the shapeless objective of satisfying users' "requirements." The assignment to particular agencies of responsibility for urgent projects is difficult; it must be done outside of normal community channels when it is done at all. Failure to assign responsibility results in duplicative efforts on the part of every agency having some interest in the project; examples of this can be cited and documented.

Security barriers prevent personnel in one agency from acquainting themselves with R&D going on in another. Another lamentably frequent occurrence is failure to set up any criteria against which to judge when a particular R&D effort has been pursued far enough and should either be abandoned or declared satisfactory.

The recommended USIB objectives and policies should be neither so broad as to lose meaning nor so narrow as to be less than comprehensive in aggregate. To be useful, they should provide for making the best possible use of community R&D laboratories, facilities,

funds, and manpower; they should authorize and encourage interagency communication and coordination; they should make the most of resources and results external to the community; they should require interchange between the community and other government agencies and between the intelligence and scientific communities; they must support federal objectives; and they should provide for measuring their own impact on community requirements and individual agency resources. Without such policy and objectives, the continuing development of more expensive equipment and more complex and intellectually demanding technology will consume more and more of the community's resources, even without unjustified duplication among the uncoordinated agencies.

As technology and R&D in data handling become more expensive, in both talent and funding, the last ounce of usefulness should be realized from every project. To this end IDH/R&D personnel should be better informed about completed and current R&D efforts everywhere. As a rough estimate, one tenth of one percent of the intelligence funds earmarked for data-handling R&D in FY 66, if spent on improvement in the information usage patterns of IDH/R&D officers, would give each of them throughout the community the equivalent of a full semester of college-level education during the year. The improvement in the resultant R&D effort would conservatively be worth 100 to 1000 times that expenditure. The distribution of the recommended listing of information services and encouragement to use them is at least a slight first step toward such self-improvement.

Technical Considerations

In its effort to identify discrete areas of intelligence data handling so as to relate the R&D to managerial responsibilities, to applications, to intelligence products, and to funding, the task team after a great deal of deliberation chose two approaches. The first of these was to classify data-handling R&D by application, and twenty-two types of application were enumerated. These range from common ones like calculation of movements (say trajectories), cryptanalysis, and document retrieval to some that may not be obvious—the monitoring of systems (say lie detection systems), image interpretation, pattern recognition, predictive calculations (say in estimates), planning (say of penetration operations), problem solving (say in inductive intelligence analysis), etc. This listing provided a basis for assessing current efforts and deficiencies.

The second approach was intended primarily to highlight ways in which R&D in data handling could improve intelligence production and management in the community. The team believes that such R&D—comprising the development of theories, advanced techniques, and equipment and the application of these to the subject in question—should have the following aims:

Experimentation with and evaluation of existing data-handling systems.
Development of criteria and measures for evaluating or designing data-handling systems.
Improvement of management procedures for allocating resources in the intelligence community.
Analysis of practices used in exploiting data and data sources.
Development and evaluation of information- or document-handling systems.
Improvement of techniques for producing and evaluating finished intelligence (e.g., quality control).
Development of validity criteria for information, including criteria for data purging.
Development of improved procedures for intelligence training (e.g., programed instruction).
Development of reporting mechanisms for R&D project funding and managerial data.

This listing served to emphasize the findings of the task team concerning the intelligence community's use of R&D resources in the data-handling area. There is no question but that this gross functional area is too large and diverse to be managed effectively as an entity. As its many constituent parts become more sharply defined, it undoubtedly will and should be split up so as to become more manageable. But the greater problem at the moment appears rather to be that a number of its essential sub-areas requiring improvement are being neglected; they are not thought of as being a part of intelligence data handling because of the parochial and limited view taken by many toward this R&D area. It was concern over this danger that prompted the above listing.

Another feature of R&D in intelligence data handling brought out by the listing is the unmistakable way it transcends the responsibilities and missions of individual agencies and so is indeed a USIB community concern. Almost without exception the intelligence pro-

duced by a given agency forms merely one part of the required community product. Thus the R&D projects of a given agency need with a few exceptions to recognize related requirements in other agencies; community priorities rather than individual agency priorities for R&D are needed; and a USIB mechanism is needed to coordinate and assign R&D responsibilities among the agencies.

In the course of the research which led to these generalizations the team uncovered certain R&D areas which seemed to demand increased attention under any criteria that might be established. The most critical single one of these is discussed below.

Indications Intelligence

Efforts to improve the processing of current intelligence information for purposes of indications and warning have been under way since 1959. Although millions of dollars and hundreds of man-years have been expended in applying automatic data processing to this effort, the results to date have been disappointing. Because of this, the level of funding for R&D in this field is currently low.

In analyzing the reasons for the past failures, it should be recognized that indications is one of the most difficult fields in all of intelligence processing. It is characterized by extremely high volumes of data, a tendency for input data to be fragmentary, redundant, and of unknown validity, wide variety in types of data, a dependence on all types of collection, severe time restrictions on processing, a critical importance for random and rare events, a tendency toward rapid changes in focus of attention, and heavy dependence on predictive evaluations. But the inherent complexity of the processing problem has been aggravated by the inadequacy of preliminary studies undertaken prior to system design. The intelligence objectives have usually been stated in such broad terms as to be practically useless to the system designer. The designer has usually been versed in some specific technology but not in intelligence. The intelligence analysts have known very little about current technology and have been too hard pressed keeping up with their work to give enough time to the system designer. The designers have concentrated heavily on statistical techniques, particularly with reference to level of military activity, which tend to obscure important anomalies rather than highlight them.

The warning problem is more a problem of logical inference and association than of statistics. Evaluations depend principally on the

talents of the analyst—his inventiveness and imagination, his ability to sense a pattern quickly, his inductive reasoning, in short, his intelligence. R&D in this field must deal, therefore, with human factors to a large degree. While the human-factors area is one in which there is much to be accomplished, it has been found to be one in which accomplishment is most difficult. A spate of human-factors work in recent years has led to a considerable amount of disillusionment. Nevertheless, several topics in this area do offer promise and should be considered in an R&D approach:

Data presentation. Many automated techniques can be used to ease the burden of the analyst in having to handle large volumes of data rapidly. Information can be so entered into a store that cumulative data on any particular topic is immediately available. Large amounts of information can be presented in simplified structure by automatic arrangement into graphic form. Different fields of information can be compared directly in combined displays and overlays. Time comparisons can similarly be made.

Time compression. Trends which may be too subtle for the analyst to note with the normal passage of time can be amplified by time-compression techniques to the threshold of recognition. Chronologically successive displays can be viewed in greatly accelerated time, and this process can be repeated (or reversed) at the analyst's desire. Such techniques can be programed on a computer if the data stores are properly arranged. In addition to bringing out past trends, they might be useful in suggesting future trends, much like extrapolating a graph beyond its plotted positions.

Query languages. Not only have computers become more powerful and economically available in recent years, the methods of utilizing them have been greatly simplified. The recent advances in query languages make the computer accessible for immediate intercommunication with the operator, so that it serves him as a direct adjunct and tool. It is now possible, therefore, for an analyst to game a complicated problem in fragmentary data analysis in which many probabilistic variations have to be considered.

Communication fundamentals. The process of communication among people involves far more than the simple transfer of information. To be considered fully successful it must create a chain or network of understanding. The physical sciences have mastered the transfer of information. The behavioral sciences have not been so successful in the communication of understanding. This is a critical

handicap for the analyst seeking to establish meaning from fragmentary information. R&D studies in the behavioral sciences might answer questions like these: Would the establishment of common goals improve the quality of communication in the warning process? Would personal contact improve understanding among the people concerned? Would group activity among analysts (like "brain-storming") heighten their imagination and contribute to solutions?

If one were to attempt now to design an ADP system to assist the indications analyst, the following methodological avenues would have to be explored:

Document search.
Interrogation of intelligence analysts.
Observation of current manual analytic processes.
Experience with previous automatic systems.
Research on types of indicators.
Analysis of the intelligence infrastructure supporting the system.
Manipulation of the ADP system under laboratory conditions.

Previous systems designers have confined their efforts almost exclusively to the first four methods, probably because these are generally straightforward and the least costly. Inadequate effort has been expended on indicator research and the infrastructure, virtually none on laboratory trials. As a result, there has been little more than an attempt to automate some part of what is already being done manually. The value of such an effort is highly questionable, given the inherent superiority of the human mind over machine capability in such areas as judgment, imagination, and inductive reasoning.

With reference to indicators, lengthy lists have been prepared by various intelligence organizations, some officially adopted by the USIB. The individual indicators are identified as ominous events or conditions which it is assumed would occur prior to hostilities. But until recently very little effort was expended on developing lists of specific phenomena that particular collectors should look for as evidence that these ominous events or conditions are taking place. Thus what is needed is extensive research to list such indicators of the indicators, the assignment of individual items on such lists to appropriate collection activities, a reporting system designed for rapid communication and processing, and extensive collation of the results in the respective indications centers.

The success or failure of any automated system is heavily dependent on the related intelligence infrastructure—coding systems, field formats, communications systems (including digital data links), interface between intelligence organizations, etc. Yet system designers working in the field of current intelligence have traditionally focused their attention on information within a particular intelligence organization treated in isolation. The result has been that, on the one hand, only a part of the data available on any given subject ever enters the system, and on the other hand, the effort required to convert into machinable form information available from other organizations swamps the personnel assigned to the task. Only by treating a given subject area, say Cuban ground forces order of battle, in its totality can an effective ADP system be developed for it.

All too little effort has been expended in attempting to analyze in depth the methods of analysis now utilized in current intelligence. It is unlikely that this can be done in the operational environment of an indications center because research and development cannot be permitted to interfere with its regular day-to-day work. What is needed, therefore, is a testing of analysis techniques using live information in a separate facility, as it were a laboratory, and comparison of the results with the regular product of the indications center. In this manner some of the areas which today are considered so difficult but which seem to offer great potential, such as cross correlation of different subject files (e.g., personnel movements with missile tests), can be explored in depth.

In summary, although the results of past efforts to improve processing capabilities in indications intelligence have been disappointing, the task is not impossible and general lines of approach can be drawn. It is believed that the present range of manual analysis can be extended significantly through these new approaches.

For Immediate Action

The task team selected five of its recommendations as in its view requiring immediate action. These are listed below. They are not in order of importance or urgency; the team believes they should all be adopted. They are not interdependent, however, and any one or any combination of them could stand alone.

The establishment of an R&D policy body to advise the Chairman, USIB. Without such a body there is nothing to which the

accomplishment of R&D projects in intelligence data handling can be related or addressed.

Improvement of indications intelligence. Early expensive failures in this field have no doubt resulted in burnt fingers; but both techniques and equipment have since improved, and new efforts should be undertaken.

Heightened use of information services. This can be promoted by compiling and disseminating a report on all information services of use to IDH/R&D officers and how to use them. Then a survey of usage patterns should follow.

The establishment of criteria for evaluating data-handling systems. For this purpose two different types of systems now in use might be singled out for systematic experimentation and evaluation. In the past, IDH systems have been developed to meet recognized requirements but without benefit of authoritative criteria to insure that the requirement was fulfilled. It is anticipated that the development of such criteria will be costly in both people and resources, but the cost is justified in view of the extremely large amounts that have been spent on unsuccessful developments.

The establishment of a feedback mechanism from consumers of intelligence to producers and IDH/R&D personnel. No mechanism exists to measure and make known the results of good or bad usage of existing data-handling capabilities in the production of finished intelligence. Results are written up and utilized by USIB with very little if any follow-up to assess why an estimate proved to be incomplete or in error. An informal review group is recommended which would determine whether or not the IDH capabilities were adequately exploited and all the available data properly used. This would encourage producers of finished intelligence to work more closely with IDH/R&D personnel. An initial analysis to determine feasible feedback techniques would require about one year.